

Xylella fastidiosa genetic traits associated to climate and host adaptation identified by population biology and genomics

Deliverable 1.2

28 December 2025

Author(s): Leonardo de la Fuente (P9- AU), Blanca B. Landa (P1-CSIC), Noé Arroyo-Vélez and Jessica Dittmer (P8-INRAE)

Funded by the European Union

DELIVERABLE INFORMATION

Start of the project:	01 Sep 2022
Duration:	48 months
Coordinator:	Blanca B. Landa
Deliverable title:	*Xylella fastidiosa* genetic traits associated to climate and host adaptation identified by population biology and genomics
Deliverable n°:	1.2
Nature of the deliverable:	Other
Dissemination level:	Public
WP contributing:	WP1
Lead beneficiary:	CSIC
Due date of deliverable:	31 Dec 2025
Actual submission date:	28 Dec 2025

Deliverable status

Ver.	Status	Date	Author(s)
1.0	Final	28 December 2025	Leonardo de la Fuente (P9- AU), Blanca B. Landa (P1-CSIC), Noé Arroyo-Vélez and Jessica Dittmer (P8-INRAE)

Disclaimer

The content of this deliverable does not necessarily reflect the official opinions of the European Commission or other institutions of the European Union.

Funded by the European Union

Table of contents

Background 4

1. Distribution and loss of conjugative plasmids in *Xylella fastidiosa* 5
2. Prophage diversity among *Xylella fastidiosa* 6
   - *2.1. Prophage diversity among Xylella fastidiosa subspecies* *6*
   - *2.2. Prophage diversity among French Xylella fastidiosa subsp. multiplex strains* *6*
3. Pseudogene content and host specialization in *Xylella fastidiosa* 7
4. Variation in natural competence across *Xylella fastidiosa* genotypes 8
5. Lack of evolutionary parallelism in host shifts of *Xylella fastidiosa* subsp. *pauca* to olive 9
6. Genomic signatures associated with host shifts of *Xylella fastidiosa* from coffee to grapevine 10
7. List of publications 10

Funded by the European Union

BACKGROUND

Xylella fastidiosa was the first plant-associated bacterium to have its genome fully sequenced in the year 2000. Obtaining good quality and complete genomes of *X. fastidiosa* strains is essential to better understand its mechanisms of virulence, host interactions, and evolutionary history. This information is extremely important to pave the way for the development of more targeted and effective strategies for diagnosis, disease management, and controlling the diseases caused by this bacterium.

Xylella fastidiosa has an average genome size of approximately 2.5 Mb, characterized by a high number of repetitive regions. These regions can include mobile elements, duplicated sequences, and other genetic elements that complicate genome sequencing and assembly when using Illumina sequencing alone. The challenges presented by the presence of repetitive regions in the genome of *X. fastidiosa* necessitate the use of complementary sequencing technologies, such as long-read sequencing platforms like PacBio or Oxford Nanopore, that can allow to achieve more complete and accurate genome assemblies. When the project BeXyl started in 2022 there were about ~140 *X. fastidiosa* whole genomes available at NCBI. During this project we have generated 315 genome sequences, increasing threefold the information available to the scientific community, including novel genomes as well as those whose quality has been improved by using long-read complementary platforms for sequencing. For a list of whole genome sequences generated during BeXyl, see Table 1-D1.2. Noteworthy, we have generated 130 genomes from strains isolated in Europe, which constitute a very valuable resource towards understanding the characteristics of the outbreaks ongoing in this continent.

This deliverable is organised into five sections addressing complementary aspects of genome evolution, host adaptation and adaptive potential in *Xylella fastidiosa*. **Section 1, "Distribution and loss of conjugative plasmids in *Xylella fastidiosa*,"** analyses the presence, absence and apparent loss of conjugative plasmids across strains, with particular emphasis on subspecies *multiplex* isolates from France and their evolutionary and epidemiological implications. **Section 2, "Prophage diversity among *Xylella fastidiosa*,"** examines the diversity, distribution and potential adaptive role of prophages across subspecies and geographic populations, including a detailed analysis of French *multiplex* strains. **Section 3, "Pseudogene content and host specialization in *Xylella fastidiosa*,"** explores patterns of pseudogene accumulation across subspecies and strains and their association with host range breadth and specialization. **Section 4, "Variation in natural competence across *Xylella fastidiosa* genotypes,"** investigates differences in recombination ability among strains from diverse subspecies and origins, highlighting implications for horizontal gene transfer and environmental adaptation. **Section 5, "Lack of evolutionary parallelism in host shifts of *Xylella fastidiosa* subsp. *pauca* to olive,"** evaluates whether independent olive-infecting populations show convergent genomic changes, revealing limited evidence for molecular parallelism. Finally, **Section 6, "Genomic signatures associated with host shifts of *Xylella fastidiosa* from coffee to grapevine,"** presents candidate genes potentially involved in the emergence of Pierce's disease of grapevine, providing hypotheses for future experimental validation.

Funded by the European Union

1. DISTRIBUTION AND LOSS OF CONJUGATIVE PLASMIDS IN *Xylella fastidiosa*

Xylella fastidiosa has a high adaptive capacity, as evidenced by its high genetic diversity, frequent recombination events and its expanded plant host range after introductions to new geographic areas, notably Europe. Extrachromosomal genetic elements such as plasmids can also play a key role for this adaptive capacity, since they can mediate the exchange of genetic material between bacteria colonizing the same habitats. The wealth of complete and high-quality *X. fastidiosa* genome sequences obtained in the last years (incl. during the BeXyl project) finally enables us to analyze plasmids associated with a variety of strains, including those responsible for new outbreaks.

In this project, we completed a genome catalogue for 75 strains belonging to the subspecies *multiplex* associated with various host plants in France. In addition, the genomes of 10 strains from the subspecies *fastidiosa* and three of *pauca*, four of *multiplex*, and two of *sandyi* isolated in the USA, Spain, Italy, Argentina, Brazil and Israel, and two strains intercepted in France or The Netherlands, were (re)sequenced and closed. All genome sequences were obtained using PacBio or Oxford Nanopore long-read sequencing technologies (for a complete list see Table 1-D1.2). A list and short description of the plasmids detected in various *X. fastidiosa* genomes is provided in Table 2-D1.2.

Among the 75 strains isolated in France, only 8 possessed a plasmid (Table 2-D1.2). This plasmid was 38 Kb in size and contained 40 genes representing the complete conjugative machinery (*tra/trb*). It was identical to the one observed in numerous ST1 strains (subspecies *fastidiosa*) and a few *multiplex* strains from the USA (strains Dixon, Riv5) (Rogers & Stenger 2012; https://doi.org/10.1371/journal.pone.0052131). Interestingly, the 8 strains harbouring the plasmid belonged to the two predominant sequence types present in France but had all been isolated in the same geographic area in the PACA region. Not a single strain from the island of Corsica or the Occitanie region possessed the plasmid. This is of interest since Dupas et al. 2023; https://doi.org/10.1038/s42003-023-04499-6) suggested that *X. fastidiosa* had likely been first introduced to the PACA region and only later to the island of Corsica. This suggests that the strains introduced to France initially had this plasmid but most strains lost it during the subsequent spread to different areas. It is also noteworthy that the ST1 strain B3_SN 17-2 from the outbreak in Israel also does not possess this plasmid, although the majority of ST1 strains sequenced from California and the Balearic Islands have it. This opens new questions regarding the fitness advantages conferred by this and other plasmids under different environmental conditions and/or in different host plants.

Funded by the European Union

2. PROPHAGE DIVERSITY AMONG *Xylella fastidiosa*

Bacteriophages play a central role in bacterial evolution by promoting gene acquisition, modification and loss, often conferring advantageous traits that enhance host fitness (Fortier & Sekulovic 2013; https://doi.org/10.4161/viru.24498). In *Xylella fastidiosa*, prophages encode genes involved in key biological functions, including stress tolerance, biofilm formation, and virulence (Jiménez et al. 2024; https://doi.org/10.1111/ppa.13992).

To assess how prophages contribute to the genetic diversity and evolutionary dynamics of *Xylella fastidiosa*, we first conducted a comparative analysis of prophage content across strains representing the full taxonomic and genetic breadth of the species. Following this species-wide assessment, we then focused on a recently established and epidemiologically relevant system to examine prophage diversity at a finer population scale. In particular, we investigated *Xylella fastidiosa* subsp. *multiplex* populations introduced into France, where limited time since introduction and restricted sequence type diversity provide a unique opportunity to explore the role of prophages in early adaptation and spread.

2.1. Prophage diversity among *Xylella fastidiosa* subspecies

Initially, as a first step we analyzed the diversity and distribution of prophages across 89 *X. fastidiosa* strains representing the three main subspecies: *fastidiosa*, *multiplex* and *pauca*, as well as the two proposed subspecies *sandyi* and *morus,* representative of 28 sequence types (ST). A total of 410 intact prophages were identified, with a notable prevalence in strains of the subspecies *sandyi* and *multiplex*. Comparative analyses revealed 105 unique prophages (Table 3A-D1.2), highlighting their role in enhancing genetic diversity through horizontal gene transfer and recombination. While some prophages were strain-specific, others were shared across multiple strains of the same subspecies or ST, suggesting clonal propagation. Genomic and phylogenetic comparisons showed clear distinctions between prophages and lytic phages and highlighted similarities among prophages from different subspecies, reflecting shared evolutionary processes. These findings will support future studies on the functional roles of specific prophage genes, their contributions to *X. fastidiosa* virulence and host range, and their potential applications in phage-based biocontrol.

2.2. Prophage diversity among French *Xylella fastidiosa* subsp. *multiplex* strains

Xylella fastidiosa was recently introduced to France, where four sequence types (STs)—ST6, ST7, ST88, and ST89—have been identified to date in three geographic regions: Corsica, PACA and Occitanie (Cunty et al. 2022, https://doi.org/10.1007/s10658-022-02492-z; Dupas et al. 2023, https://doi.org/10.1038/s42003-023-04499-6). As recently stablished populations, their adaptive potential and genetic plasticity are critical to colonize and persist in novel

Funded by the European Union

ecological niches. Given the fundamental role of prophages in bacterial adaptation, in this work, we investigated their frequency and diversity in *X. fastidiosa* subsp. *multiplex* strains isolated in France.

We analysed a dataset of 67 complete *X. fastidiosa* subsp. *multiplex* genomes representing all STs reported in France to date, many of which were generated within the framework of the BeXyl project. Using VIBRANT for prophage detection and CheckV for validation, we identified a total of 517 prophage sequences (Table 3B-D1.2). Strains harbored between 5 and 10 prophages, with a median of seven per genome, representing a relatively high prophage load compared to other phytopathogens with similar lifestyles (Evseev et al. 2023, https://doi.org/10.3390/ijms24021586).

The prophage repertoire of French *multiplex* strains was highly diverse and could be classified into 37 viral species clusters using VIRIDIC. Taxonomic assignment with PhaGCN classified 20 clusters within the class *Caudoviricetes*, comprising double-stranded DNA tailed phages, while 17 clusters were assigned to the genus *Lophivirus*, a group of single-stranded DNA filamentous phages. Both the number and composition of prophages varied across sequence types, with ST6 strains exhibiting the highest prophage burden.

Marked differences in prophage repertoires were also observed between geographic populations, whereas strains sharing the same ST and region displayed strong intra-population homogeneity. This pattern supports the hypothesis that each French *multiplex* population originated from independent introduction events, as previously proposed (Dupas et al. 2023, https://doi.org/10.1038/s42003-023-04499-6; Cunty et al. 2025, https://doi.org/10.1016/j.dib.2025.111816). Notably, one prophage species cluster was conserved across all analyzed strains, consistent with an ancestral acquisition that likely conferred selective advantages and became fixed within the ancestral population. This conserved prophage therefore represents a strong candidate for future functional characterization. In contrast, several prophage species clusters were population-specific; these accessory elements likely reflect local selective pressures and may underlie region-specific adaptive strategies, providing valuable insights into the evolutionary trajectories and dispersal potential of *X. fastidiosa* populations in France.

3. PSEUDOGENE CONTENT AND HOST SPECIALIZATION IN *Xylella fastidiosa*

Pseudogenes are regarded as “junk” DNA, representing vestigial functions no longer needed for fitness. Accordingly, higher number of pseudogenes in a bacterial human pathogen were previously hypothesized to be a hallmark of host specialists.

Funded by the European Union

In the study by Kaur et al., (2025) (https://doi.org/10.1128/aem.02070-24), we tested this hypothesis on the emergent bacterial plant pathogen *Xylella fastidiosa* linking pseudogene makeup and host specificity. *Xylella fastidiosa* is an ideal subject for these studies by being a xylem-limited pathogen that underwent extensive genome reduction. Using natural host range data of 151 strains and Pseudofinder analysis on *X. fastidiosa* whole genome sequences, we observed that *X. fastidiosa* subsp. *sandyi* had the highest pseudogene content followed by subsp. *morus*, while subsp. *pauca*, *fastidiosa* and *multiplex* had the lowest. The first two subspecies are known to have a limited host range compared to the others, aligning with the hypothesis of a greater number of pseudogenes corresponding to narrower host range. Based on our pseudogene analysis, we propose that strains from subsp. *pauca*, *fastidiosa* and *multiplex* could be considered as generalists, therefore posing a greater risk of host jumps. On the other hand, subsp. *sandyi* and *morus* could be considered specialists, posing less risk of expansion to different hosts. Interestingly, weed isolates are presumably host specialists since they had the highest pseudogene content. Data on pseudogene content in strains isolated in Europe and elsewhere are shown in Table 4-D1.2. Here we generated hypotheses on the host range of *X. fastidiosa* strains that need to be tested experimentally to help understand this devastating plant pathogen.

4. VARIATION IN NATURAL COMPETENCE ACROSS *Xylella fastidiosa* GENOTYPES

Natural competence is one of the mechanisms of horizontal gene transfer, an important process that contributes to host-use evolution and other types of environmental adaptation in bacteria. Recently, the plant pathogen *Xylella fastidiosa* has undergone expansion of its host and geographic ranges. Natural competence has been empirically documented for a few strains of *X. fastidiosa*, but its prevalence across genotypes and populations is largely unknown. In the study by Liu et al., (2025) (https://doi.org/10.1371/journal.ppat.1013757), we characterized *in vitro* the natural competence of 142 *X. fastidiosa* strains from diverse hosts and geographic origins, and revealed substantial variability among strains, particularly across subspecies. *X. fastidiosa* subsp. *fastidiosa* strains were largely naturally competent, while only 15% of studied subsp. *multiplex* strains showed recombination, and none of the strains classified in other subspecies were competent. While recombination rates in vitro were associated with subspecies classification, host and climatic variables from the area of isolation did not explain differences in recombination across strains. Overall, this study highlights the variability of natural competence among *X. fastidiosa* strains, that could have an impact on their potential for adaptation to the environment. Data on recombination rates of strains isolated in Europe and elsewhere are shown in Table 5-D1.2.

Funded by the European Union

5. LACK OF EVOLUTIONARY PARALLELISM IN HOST SHIFTS OF *Xylella fastidiosa* subsp. *pauca* TO OLIVE

The genetic and evolutionary history of the emergence of *X. fastidiosa* in olive trees in Brazil, as well as molecular parallelism between host shifts to olives in ST16 (Brazil) and ST53 (Italy) was investigated in the study by Donegan et al. (2023) (https://doi.org/10.1111/mpp.13316). In examining the two olive-infecting clades (Brazil and Italy) for genetic convergence, we found that neither gene gains and losses nor SNPs converged between the Brazilian and Italian olive-infecting populations. Using a data set of *X. fastidiosa* subsp. *pauca* genome sequences from isolates infecting coffee and olive trees (n=39), we used TreeWAS to identify SNPs and accessory genes associated with (rather than unique to) olive strains. We repeated this technique to detect core SNPs or accessory genes associated with citrus- and coffee-infecting strains as well.

To examine potential molecular parallelism between ST16 and ST53, we aligned a set of 39 *X. fastidiosa* subsp. *pauca* strains (Table 6-D1.2). Within subsp. *pauca*, we used Scoary to identify olive-specific variation in the core and accessory genome within ST16 and ST53, that is, genes or SNPs present/absent in all olive-infecting strains relative to all sister taxa (i.e., coffee-infecting strains) within the clade. ST16 and ST53 both showed olive-specific gene gain and loss events, as well as olive-specific SNPs. In ST53, we identified core-genome SNPs in seven genes; two of the genes were excluded because of gapped alignments and two of the genes were hypothetical proteins. The three remaining genes (*clp*, *rpfC*, and *ilvG*) each contained a single nonsynonymous mutation. None of these mutations occurred in inferred recombinant regions. In the ST53 accessory genome, 96 gene loss and six gene gain events were unique to olive. Of the 96 gene loss events, there were 15 genes involved in toxin-antitoxin systems, many of which were on a large plasmid deletion (Table 6-D1.2). Additionally, seven genes were annotated as Type IV secretion system proteins, as well as 18 phage-related accessory gene losses. Neither genes with SNPs nor gene gain and loss events were shared between the two clades, that is, there were no genes with convergent mutations among olive-specific variants in ST53 and ST16.

Additionally, using the same set of *X. fastidiosa* subsp. *pauca* genomes ($n = 39$), we used TreeWAS to identify SNPs and accessory genes associated with (rather than unique to) olive strains. Neither core SNPs nor accessory genes were associated with coffee strains or citrus strains, and no core SNPs were associated with olive strains. One gene loss (hypothetical protein: group_1969) was significantly associated with olive-infecting strains. While group_1969 was absent from all olive strains from ST53, it was present in some olive strains within ST16. A BLAST search revealed that this protein was a tail fibre domain-containing protein. Although statistically significant, this gene loss is probably not necessary for *X. fastidiosa* adaptation to olive due to its presence in ST16.

Funded by the European Union

6. GENOMIC SIGNATURES ASSOCIATED WITH HOST SHIFTS OF *Xylella fastidiosa* FROM COFFEE TO GRAPEVINE

When pathogens are repeatedly introduced into new environments, host jumps may occur into naïve taxa. Given the magnitude of the global plant trade, this process can lead to frequent disease emergence as interactions between previously isolated pathogens and new plant hosts become possible. *Xylella fastidiosa* is a recurring nuisance. This bacterial pathogen has recently emerged in novel geographic locations infecting a breadth of host plants. An introduction of *X. fastidiosa* subsp. *fastidiosa* from Central America to the USA several hundred years ago has since been the source of outbreaks across the globe. In the USA, particularly in California, the introduced bacterium is frequently found in European grapevine (*Vitis vinifera*).

In the study developed in Kahn et al. (2025) (https://doi.org/10.1099/mgen.0.001447) (Table 7-D1.2) a suite of genes that are related to an *X. fastidiosa* host switch to European grapevines was identified. Our findings support the hypothesis that the introduction event, from Central to North America leading to the emergence of Pierce's disease of grapevines, may have involved a shift in host use, potentially accompanied by reduced infectivity towards the ancestral host and improved compatibility with the novel host. The gene lists shown in Table 7-D1.2, however, should be interpreted as hypothesis to be tested at this stage.

7. LIST OF PUBLICATIONS

- Amoia SS, Saponari M, Ligorio A, Saldarelli P, Boscia D, Giampetruzzi A. (2025). Complete genome assembly of *Xylella fastidiosa* subsp. *fastidiosa* (ST1) strains "BA_GR31" and "BA_AL21" associated with a new outbreak in Apulia, Italy. Journal of Plant Pathology 1-4.
- Arroyo-Velez N, Jacques M-A, Dittmer J. Prophage diversity among French *Xylella fastidiosa* subsp. *multiplex* strains. Manuscript in preparation.
- Briand M, Jacques M-A, Dittmer J (2025) The hidden life of *Xylella*: mining the NCBI Sequence Read Archive reveals potential new species, host plants, and infected areas for this elusive bacterial plant pathogen. Applied and Environmental Microbiology 91: e00913-25
- Dittmer J, Vo Hoang Quoc T, Briand M, Legendre B, Cunty A, Boutigny A-L, Bahut M, Cesbron S, Jacques M-A. 2025. Comparative genomics of *Xylella fastidiosa* subsp. *multiplex* strains from France: Genetic diversity and pathogen dynamics. *Manuscript in preparation.*
- Donegan MA, Coletta-Filho HD, Almeida RPP. Parallel host shifts in a bacterial plant pathogen suggest independent genetic solutions. Mol. Plant Pathol. 2023 Jun;24(6):527-535. doi: 10.1111/mpp.13316.

Funded by the European Union

- Fernandez F D, Arias-Giraldo LF, Tolocka PA, Guzman FA, Olivares García C, Roca ME, Otero M L, Conci LR, Landa BB, Haelterman RM. 2024. Phylogenomic analysis of *Xylella fastidiosa* subsp. *pauca* strains from olive and almond trees in Argentina. European Journal of Plant Pathology 171: 593–602. doi: 10.1007/s10658-024-02969-z
- Kahn AK, Cervantes J, Almeida RPP. A pathogen of good taste: genetics of a bacterial host jump of the plant pathogen *Xylella fastidiosa* from coffee to wine grapes. Microb Genom. 2025 Jul;11(7):001447. doi: 10.1099/mgen.0.001447.
- Kaur N, Potnis N, De La Fuente L. 2025.Pseudogenes and host specialization in the emergent bacterial plant pathogen Xylella fastidiosa. Appl. Environ. Microbiol. 91:e02070-24. doi: 10.1128/aem.02070-24
- Liu R, MP Velasco, LF Giraldo, M Donegan, N Hardy, N Potnis, RPP Almeida, BB Landa, L De La Fuente. Natural competence in the bacterial pathogen *Xylella fastidiosa* varies across genotypes and is associated with adhesins. PLoS Pathogens 21(12):e1013757. doi: 10.1371/journal.ppat.1013757
- Morán, Barbé, Selles-Ribera, Monterde, Marco-Noales. 2025. Tracing the origin of the new *Xylella fastidiosa* subsp. *fastidiosa* outbreak in Extremadura (Spain), through comparative genomic analysis and phage typing. *Manuscript in preparation*
- Shantharaj, De La Fuente. Isolates of *Xylella fastidiosa* obtained from blueberry and grape plants in Alabama. *Manuscript in preparation*
- Velasco-Amo MP, LF Arias-Giraldo, M Román-Écija, L De La Fuente, E Marco-Noales, E Moralejo, JA Navas-Cortés, BB Landa. 2023a. Complete circularized genome resources of seven strains of *Xylella fastidiosa* subsp. *fastidiosa* using hybrid assembly reveals unknown plasmids. Phytopathology 113:1128-1132. doi: 10.1094/PHYTO-10-22-0396-A
- Velasco-Amo MP, LF Arias-Giraldo, M Román-Écija, A Giampetruzzi, M Saponari, BB Landa. 2023b. Complete circularized genome resources for *Xylella fastidiosa* subsp. *sandyi* strains CO33 and CFBP8478. Phytopathology 110:969-972. doi: 10.1094/PHYTO-01-20-0012-A.
- Velasco-Amo M P, Arias-Giraldo LF, Landa BB. 2024. Complete genome assemblies of several *Xylella fastidiosa* subspecies *multiplex* strains reveals high phage content and novel plasmids, Phytopathol. Mediterr. 63:15–23. doi: 10.36253/phyto-14931
- Zúñiga-Pereira AM, Khan A, Donegan M, Sánchez-Arguedas V, Valerio-Quirós G, González-Pujol M, Rodríguez-Murillo N, Ramírez-Mora T, Molina-Mora JA, Almeida RPP, Chacón-Díaz C. Phenotypic and Genomic characterization of endemic *Xylella fastidiosa* subsp. *fastidiosa* from Costa Rica. *Manuscript in preparation*

Funded by the European Union